MINOS Production Database Proposal

Using Oracle on Linux
with AMD Opteron Hardware
from Sun Microsystems
Presented by Richard Jetton
for CSS-DSG on April 14, 2004

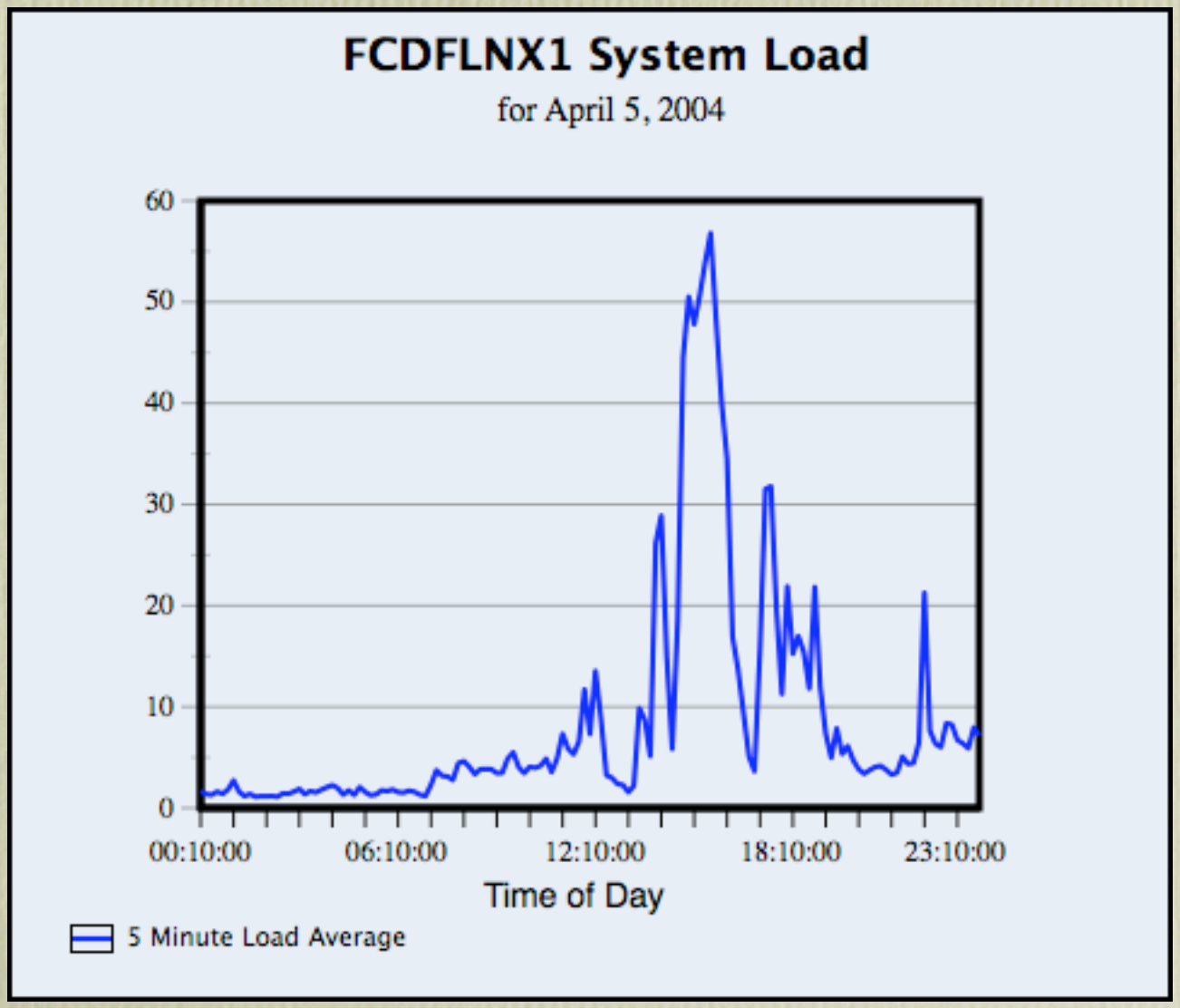
CSS-DSG Goals for MINOS

- 64 bit Oracle on 64 bit Linux
  - Larger address space compared to 32 bit systems
    - Larger SGA (breaks the 2 GB limit)
- “Tier One” hardware support
  - Comparable to that of our current database hardware
- Should be viable for 3 to 5 years
- Reliable and maintainable

Ongoing Oracle on Linux Work by DSG

- fcdflnx1 entered quasi-production use in July 2002
  - SGI 1450, Quad Pentium 3 Xeon, 4 GB RAM
  - Serves metadata for all CDF offline analysis jobs
  - Reliable, with uptime limited by kernel patches
  - User load growing, with many recent spikes
  - Database replicated from CDF online
  - Users perform only queries against this database

Twenty-four Hour Sample of Activity on fcdflnx1



Database Performance Testing using SAM DBServer

- Used SAM services installed on d0ora1
- Used “clientStation.py” to generate database load
- SQL statement timings were inserted into a MySQL database automatically
- Custom test harness written to drive a test run and gather statistics
- Test runs lasted about 4 hours, and were begun at various times of day and night (and weekends)
- Databases imported from d0ofdev instance

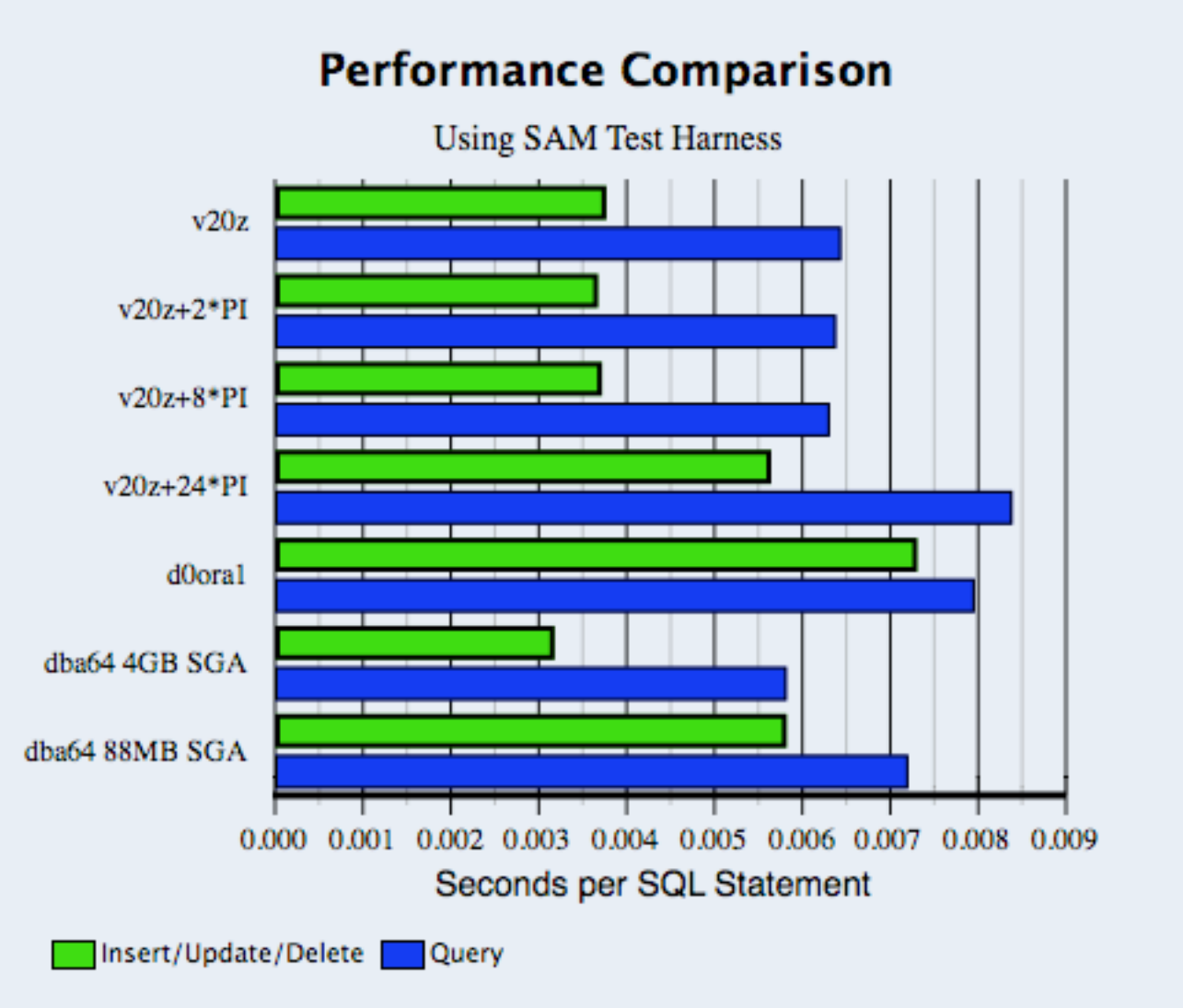
Database Computers Tested

- sunminos (aka v20z, instance name sunminos)
  - Sun v20z dual Opteron, 4 GB RAM
  - External Ultra160 SCSI drives for Oracle
  - RedHat Advanced Server 3.0, Update 1
- dba64 (instance name minosdev)
  - “White box” dual Opteron, 8 GB RAM
  - Internal Ultra320 SCSI drives for OS and Oracle
  - RedHat Advanced Server 3.0, Update 1
- d0ora1 (instance name d0ofdev)
  - Sun E4500 UltraSPARC2 8-way, 4 GB RAM
  - External SCSI drives for Oracle
  - Solaris 9

Background Load on the Database Computers

- SETI@Home
  - Used to insure few (if any) idle CPU cycles
  - Used on the v20z and dba64, only
- CHURN
  - Looping shell script to flush OS buffers and caches, and create an artificial disk I/O load
  - Used on the v20z, only
- PI
  - Looping shell wrapper for a bc script to compute 3000 digits of π
  - Used on the v20z, only

Test Results



Comments about the Results

- All tests were done using Oracle v9.2, 64 bit
- The d0ora1 result is from the “best” run, started immediately after a reboot
- All results for the v20z and dba64 include 2 running SETI@Home processes
- Generalizing from the two dba64 results:
  - Sam related database activities are strongly influenced by SGA size

Recommendations

- Purchase the following from Sun Microsystems:
  - A v20z with model 248 Opteron CPUs
  - 8 gigabytes of RAM
  - 2 36 gigabyte internal SCSI drives
  - 2 Qlogic fibre channel host bus adapters
  - 3510 fibre channel drive chassis
  - RedHat Advanced Server 3.0
  - Sun service and support, 3 year contract
- Later this year, a review should be made to determine if a similar system should be ordered for MINOS production use